

MENU

SEARCH

INDEX

DETAIL

JAPANESE

BACK

NEXT

14 / 18

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-214004
 (43)Date of publication of application : 20.08.1996

(51)Int.Cl. H04L 12/28
 H04Q 3/00

(21)Application number : 07-282902 (71)Applicant : INTERNATL BUSINESS MACH CORP <IBM>
 (22)Date of filing : 31.10.1995 (72)Inventor : CALVIGNAC JEAN
 VERPLANKEN FABRICE
 ORSATTI DANIEL
 BASSO CLAUDE

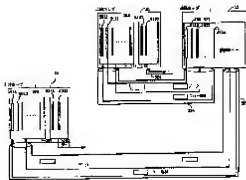
(30)Priority
 Priority number : 94 94480125 Priority date : 08.11.1994 Priority country : EP

(54) FLOW CONTROL AT EVERY HOP OF ATM NETWORK

(57)Abstract:

PROBLEM TO BE SOLVED: To make low-loss traffic control possible even when the capacity of a buffer is small by realizing fairness.

SOLUTION: A communication system contains a plurality of hops 50, 51, and 52 interconnected to each other through links 524 and 525 each constituted of a plurality of connections. The traffic among the hops 50, 51, and 52 is shared by a reserved band width service and a non-reserved band width service. The control of the non-reserved band width service is composed of back pressure mechanisms at every hop. When the traffic entering a certain hop exceeds an upper/lower threshold, the back pressure mechanism of the hop generates a start/stop back pressure primitive for restating/adjusting the input traffic. When traffic congestion occurs, the mechanism can selectively stop the connection which is a cause of the congestion without giving any influence to the other part of a link traffic or can globally stop the traffic of all links.



LEGAL STATUS

[Date of request for examination]

05.11.1997

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 2986085

[Date of registration] 01.10.1999

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

特開平8-214004

(43)公開日 平成8年(1996)8月20日

(51)Int.Cl. ⁴ H 0 4 L 12/28 H 0 4 Q 3/00	識別記号 9466-5K	庁内整理番号 H 0 4 L 11/ 20	F I G	技術表示箇所
---	-----------------	--------------------------	----------	--------

審査請求 未請求 請求項の数17 O L (全 16 頁)

(21)出願番号 特願平7-282902	(71)出願人 390009531 インターナショナル・ビジネス・マシー ズ・コーポレーション INTERNATIONAL BUSIN ESS MACHINES CORPO RATION アメリカ合衆国10504、ニューヨーク州 アーモンク (基地なし)
(22)出願日 平成7年(1995)10月31日	(72)発明者 ジーン・カルビグナク フランス06610、ラ・ゴード、チェミン・ デス・バリエレス 187
(31)優先権主張番号 9 4 4 8 0 1 2 5 . 7	(74)代理人 弁理士 合田 潔 (外2名)
(32)優先日 1994年11月8日	
(33)優先権主張国 フランス (F R)	

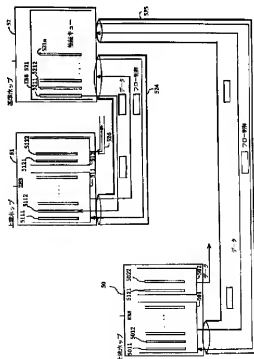
最終頁に続く

(54)【発明の名称】 ATMネットワークのホップ毎のフロー制御

(57)【要約】

【課題】 フェアネスを実現しつつ、バッファ容量が小さくても損失のないトラフィック制御を可能にする。

【解決の手段】 通信システムは、複数の接続からなるリンク524、525によって相互接続された複数のホップ50、51、52を含む。ホップ間のトラフィックは予約帯域幅サービスと非予約帯域幅サービスによって共有される。上記非予約帯域幅サービスの制御はホップ毎のバックプレッシャ・メカニズムよりなる。あるホップに入るトラフィックが上/下のしきい値を超えると、バックプレッシャ・メカニズムは、入力トラフィックを再開/調整するためにスタート/ストップ・バックプレッシャ・プリミティブを生成する。輻輳の場合、このメカニズムは、輻輳の一因になっている接続をリンク・トラフィックの他の部分に影響を与えずに選択的にストップさせるか、リンク全体のトラフィックをグローバルにストップさせることができる。



【特許請求の範囲】

【請求項1】複数の接続からなるリンクによって相互接続された複数のホップを含む通信システムに実現され、前記接続内のトラフィックはベスト・エフォート・サービスと予約帯域幅サービスに共有され、各ホップにおいて、前記ベスト・エフォート・サービスの制御は、単位接続とホップ毎のグローバル・バックプレッシャ・メカニズムの両方からなることを特徴とする、トラフィック制御装置。

【請求項2】前記ホップ毎のバックプレッシャ・メカニズムは、前記ホップに入るトラフィックが所定しきい値を超えたときに上流ホップにバックプレッシャ・プリミティブを生成する、請求項1記載のトラフィック制御装置。

【請求項3】各ホップは、サイズNの入力バッファと、出力リンク当たりサイズMの出力バッファとからなり（MとNは1より大きい整数）、いずれのバッファもベスト・エフォート・トラフィック専用であって、前記ホップに接続された接続毎にキューを提供するキューイング・エリアからなり、前記接続は、その対応するキューに置かれたトラフィック・セルがなくなるとも1つある場合はアクティブとみなされる、請求項1または2に記載のトラフィック制御装置。

【請求項4】前記入力バッファは、グローバル上しきい値 IBB_{th} とグローバル下しきい値 IBB_{tl} よりなり、 $IBB_{th} > IBB_{tl}$ であって、前記上しきい値は、前記上しきい値を超える前記入力バッファ内の余分なバッファ・エリアが、上流ホップの往復遅延合計より大きくなるようにセットされ、前記下しきい値は、上流ホップの最少往復遅延よりも大きい、請求項3記載のトラフィック制御装置。

【請求項5】前記入力バッファは、選択的上しきい値 IBB_{st} と選択的下しきい値 IBB_{sl} からなり、 $IBB_{st} > IBB_{sl}$ であって、前記しきい値は、上流ホップからのアクティブ接続の数に従って動的に調整される、請求項3または4に記載のトラフィック制御装置。

【請求項6】各ホップにおいて、前記上下の選択的しきい値 IBB_{st} 、 IBB_{sl} は、前記ホップに入るアクティブ接続の数が1と2の間にある場合、 $IBB_{st} = N/2$ 、 $IBB_{sl} = N/4$ 前記ホップに入るアクティブ接続の数が3と4の間にある場合、 $IBB_{st} = N/4$ 、 $IBB_{sl} = N/8$ 前記ホップに入るアクティブ接続の数が5と8の間にある場合、 $IBB_{st} = N/8$ 、 $IBB_{sl} = N/16$

前記ホップに入るアクティブ接続の数が9と16の間にある場合、 $IBB_{st} = N/16$ 、 $IBB_{sl} = N/32$

前記ホップに入るアクティブ接続の数が17と32の間

にある場合、 $IBB_{st} = N/32$ 、 $IBB_{sl} = N/64$

前記ホップに入るアクティブ接続の数が3より大きい場合、 $IBB_{st} = N/32$ 、 $IBB_{sl} = N/64$

というルールに従って調整される、請求項5記載のトラフィック制御装置。

【請求項7】あるホップ接続のバッファ・エリアにキューイングされたセル数が、前記動的選択的上しきい値を超え、前記ホップは、前記接続からなるリンク上で後向きに、「前記接続の選択的ストップ」情報を持つ選択的バックプレッシャを送る、請求項2乃至6に記載のトラフィック制御装置。

【請求項8】ある接続のバッファ・エリアにキューイングされたセル数が、前記動的選択的下しきい値に等しいかより小さく、前記接続が選択的バックプレッシャを受けているとき、前記ホップは、前記接続よりなるリンクで後向きに、「前記接続の選択的スタート」情報を持つ選択的バックプレッシャを送る、請求項2乃至7に記載のトラフィック制御装置。

【請求項9】あるホップのキューイング・エリアに置かれたセル数が、前記グローバル上しきい値を超え、前記ホップは、前記リンクをグローバルに停止させるために、前記ホップに入る上流リンクで後向きに、「前記入力リンクのグローバル・ストップ」情報を持つグローバル・バックプレッシャを送る、請求項2乃至8に記載のトラフィック制御装置。

【請求項10】あるホップのキューイング・エリアに置かれたセル数が、前記グローバル下しきい値に等しいかより小さく、前記ホップに入るリンクがすでにグローバル・バックプレッシャを受けている場合、前記ホップは、前記リンクをグローバルにリスタートさせるために、前記入力リンクで後向きに「前記入力リンクのグローバル・スタート」情報を持つグローバル・バックプレッシャを送る、請求項2乃至9に記載のトラフィック制御装置。

【請求項11】選択的に停止された接続は、前記接続が選択的スタート・バックプレッシャを受け取り、前記接続を含む前記リンクがグローバル・バックプレッシャを受けていない場合にのみリスタートでき、グローバルに停止されたリンクは、前記リンクがグローバル・スタート・バックプレッシャを受取った場合にのみリスタートできる、請求項2乃至10に記載のトラフィック制御装置。

【請求項12】ホップは、前記ホップに入る各リンクに関連づけられ、前記ホップが前記入力リンクにグローバル・ストップ・バックプレッシャを送った後に前記入力リンクから受信した余分なセルをカウントする手段を含み、前記入力リンクのカウント値が前記入力リンクの往復遅延より大きいか前記遅延に等しいとき、前記入力リ

3

ンクからの余分なセルは棄却され、前記ホップは前記グローバル・ストップ・バックプレッシャを前記入力リンクに再送する、請求項1乃至11に記載のトラフィック制御装置。

【請求項13】ホップは、前記ホップに入る各接続に関連づけられ、前記ホップが前記接続よりなるリンクで前記接続に選択的ストップ・バックプレッシャを送った後に前記接続から受信した余分なセルをカウントする手段を含み、前記接続のカウント値が、前記接続よりなる前記リンクの往復遅延より大きい前記遅延に等しいとき、前記接続からの余分なセルは棄却され、前記ホップは前記選択的ストップ・バックプレッシャを前記接続に再送する、請求項1乃至12に記載のトラフィック制御装置。

【請求項14】予約帯域幅の1部がベスト・エフォート・サービスの制御トラフィックに予約された、請求項1乃至13に記載のトラフィック制御装置。

【請求項15】前記バックプレッシャ情報はATM制御セルによって伝達され、前記セルは、前記ベスト・エフォート・サービスの前記制御トラフィックに予約された前記予約帯域幅の前記1部によって転送される、請求項14記載のトラフィック制御装置。

【請求項16】前記ATM制御セルは、1乃至12の選択的バックプレッシャ情報と1つのグローバル・バックプレッシャ情報を伝達できる、請求項15記載のトラフィック制御装置。

【請求項17】接続は、ATMバーチャル・サーキット・サービスのスイッチ・バーチャル・サーキットと定義されるか、ATMバーチャル・バス・サービスのバーチャル・サーキットと定義されるか、或いはATMローカル・エリア・ネットワークのエミレション・サービスにおける1組のメディア・アクセス制御(MAC)アドレス(MACソース・アドレス、MACターゲット・アドレス)と定義される、請求項1乃至16に記載のトラフィック制御装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、一般的にはネットワーク通信システムに関し、特に高速ATMネットワークのベスト・エフォート・サービス(best effort service)をサポートするフロー制御メカニズムに関する。

【0002】

【従来の技術】ATM(非同同期送モード)は次世代の高速ネットワークの基礎技術になるものである。

【0003】高速ネットワークは、トラフィックやサービス品質(QOS)に関する要求の異なる多種多様なアプリケーションをサポートする。アプリケーションがどのように多様であると、求められるフロー制御方式も異なってくる。例えばマルチメディアや時間重視型データ

4

・アプリケーション等のアプリケーションは、遅延とスループットのレベルは保証されなければならないが、損失には耐えることができる。一方、LANトラフィック等、他のアプリケーションは遅延やスループットの変動には耐えられないが、損失には非常に弱い。

【0004】ATM予約帯域幅(RB)サービスでは、ユーザはデータを転送する前に、コールのセットアップ時にネットワークとトラフィック・コントラクトを確立する必要がある(図1参照)。このコントラクトには、目的のQOSクラスとトラフィック記述子のセットの指定が含まれる。ネットワークはリソース割当てを通じて、ATM接続に目的のQOSを提供するか、コールを拒否する。通信元と通信先の間で割当てられる帯域幅は、統計的多重化ゲインを活かすためにビーク・レートよりも小さいことがあるが、これは複雑なソフトウェアを必要とし、トラフィックの輻輳やデータ損失につながることもある。

【0005】LAN環境では、通信元はバースト型で予測不可能である。トラフィックは数秒の範囲の時間スケールで変動性が極めて高い。このような予測不可能な通信元の場合、ネットワーク内の大きな損失を防ぐためにビーク・レートを割当ててもできよう。しかし、帯域幅は休止期間にも予約されるからネットワークの使用効率は悪くなる。

【0006】ネットワークのリンク使用率を高める方法には、未だにベスト・エフォート(best effort, BE)サービスと呼ばれる非予約帯域幅(NRB)サービス・クラスの追加によるものがある(図2参照)。このサービス・クラスでは、帯域幅は全く予約されず、通信元はRBトラフィックに影響を与えずに、出来る限り多くの帯域幅を獲得して、「最善の努力」で転送できる。従ってNRBサービスは、前もって割当てられるリソースはなく、通信元を制御するためにはフロー制御方式を要する。その第1の目的はネットワークの輻輳を避けることであり、第2の目的は異なる通信元のトラフィックのバランスをとることである。このような方式の1つ、バックプレッシャというコンセプトは広く研究されすでに実現されている。バックプレッシャの原理は、損失が生じる前にネットワークの輻輳ノードで入力トラフィックを止めることである。情報はバックプレッシャ信号によって運ばれる。バックプレッシャ信号は、ネットワークの上流(upstream)ノードから来たトラフィックをスタート/ストップさせる。この方式で、全ての通信元は対応するバックプレッシャ信号を受信したとき、そのNRBトラフィックをリスタート/ストップできることになっている。バックプレッシャのコンセプトについては、例えばT. L. Rodehefferの"Experience with Autonet", Computers Networks and ISDN Systems, vol. 2, 5, 1993, p. 623-629に説明されている。

【0007】前記の文書に説明されているもの等、シン

50

5

ブルなバックプレッシャ・メカニズムの欠点は、フェアネスを実現できないことである。輻輳の場合には、ノードはバックプレッシャ信号を後向きにその上流ノードに送り、輻輳の一因になっているかどうかには無関心に輻輳を広げ接続をブロックする。フェアネスを実現するにはバックプレッシャ情報は選択できなければならず、輻輳の一因になっている接続に対してのみ機能しなければならない。

【0008】このような選択的輻輳メカニズムは、B. J. VickersとT. Sudaによる"Connectionless Service for Public ATM Networks", IEEE Communications Magazine, August 1994, p. 34-42に説明されている。この文書で提案されているのは、ベスト・エフォート・サービスをホップ毎のフロー制御によってATMネットワークに実現できる方式である。ホップについては周知の通り(図3参照)であり、その端部に標準インタフェースを持つネットワーク要素と定義できる。標準インタフェースは、ユーザ・ネットワーク・インタフェース(UNI)、ネットワーク/ネットワーク・インタフェース(NNI)、または私設ネットワーク/ネットワーク・インタフェース(PNNI)である。このように定義したホップは、公衆ネットワークや私設ネットワークの全体か、ネットワーク内のATMスイッチだけである。後者の文書に述べられているメカニズムの欠点の1つは、ノードは多数の接続からなるリンクのトラフィックを調整する必要があるときは、接続毎に(1度に接続)調整する必要があることである。各接続の伝播遅延が加わるから信号が送られた後、トラフィック全体が実際にストップする前に、バッファに格納されるATMセルの数は非常に大きくなる。トラフィック損失の無いことを保証するためには、必要なバッファ・スペースは途方もなく大きくなり、効率のよいロードウェアと両立しなくなる。また格納された全てのセルを転送することになるが、その遅延も相当に長くなる。

【0009】

【発明が解決しようとする課題】本発明の目的は、フェアネスに対応し、バッファ容量が小さくても損失のないトラフィック制御装置を提供することである。

【0010】

【課題を解決するための手段】前記目的を達成するため、通信システムにおいてトラフィックを制御する装置が本発明に従って提供される。通信システムはリンクによって相互接続された複数のホップからなる。各リンクは、予約帯域幅サービスとベスト・エフォート・サービスとで共有される複数の接続からなる。ベスト・エフォート・サービスは、接続単位でもグローバルなホップ毎のバックプレッシャ・メカニズムでもサポートされる。バックプレッシャ・メカニズムは2つのプリミティブに対応する。どのホップでも1接続を制御できる選択的バックプレッシャ・プリミティブと、グローバル輻輳の場合

6

合には、どのホップでもベスト・エフォート接続の場合のように多くの選択的バックプレッシャ・プリミティブを送らずに、リンクを制御でき、トラフィックをリンク・レベルでブロックするグローバル・バックプレッシャ・プリミティブである。特に本発明は、リンクによって相互接続された複数のホップからなる通信システムにおいて実現されるトラフィック制御装置を含み、前記リンクは、複数の接続内の前記トラフィックがベスト・エフォート・サービスと予約帯域幅サービスとで共有される前記複数の接続からなり、各ホップにおいて、前記ベスト・エフォート・サービスの制御は、単位接続と、グローバルなホップ毎のバックプレッシャ・メカニズムとの両方からなることを特徴とする。各ホップはホップに接続されたNRBライン毎にキューを提供するキューイング・エリアからなる。接続(ライン)は、接続されたそのキューに少なくとも1つのトラフィック・セルがある場合にはアクティブとみなされる。バックプレッシャ・プリミティブは、ホップに入るトラフィックが所定の選択的グローバルしきい値を超えるとき、ホップによって生成される。これら選択的しきい値は、上流ホップからのアクティブ接続数に従って動的に調整される。

【0011】本発明の他の目的は、輻輳したリンクや接続でトラフィック制御情報をブロックされることのない、効率のよいトラフィック制御装置を提供することである。これは、予約帯域幅サービスの1部をベスト・エフォート・サービスのトラフィック制御のために予約し、前記ベスト・エフォート・サービスのトラフィック制御セルを前記予約された部分で転送することによって達成される。

【0012】本発明の他の目的は、出力するトラフィック制御が少ないトラフィック制御装置を提供することである。これは、1乃至12の選択的制御情報と、1つのグローバル制御情報を前記トラフィック制御セルのそれぞれに組込むことによって達成される。

【0013】

【発明の実施の形態】図4は、ホップ毎のフロー制御が実現される6ホップ構成を示している。ホップS0、S1、...、S5はそれぞれ、データ・トラフィックの送信先のホップからフロー制御情報を受信でき(前記ホップは、前記受信側ホップを基準として下流のホップと定義される)、またデータ・トラフィックの受信先のホップにフロー制御情報を送ることができる(前記ホップは、前記送信側ホップを基準として上流のホップと定義される)。例えばS1を基準ホップとすると、S0はS1の上流ホップ、S2とS5はS1の下流ホップである。

【0014】図5は図4の詳細図で、3つの相互接続ホップ50、51、52が相互にどのように動作しているかの概略を示している。ホップ50、51からホップ52へのデータ・トラフィックと、ホップ52からホップ50、51への制御トラフィックのみを示しているが、A

7
 TMは全2重なの、逆のパス、つまりホップ52からホップ50、51へのデータ・トラフィックと関連するフロー制御も実現される。本発明の好適な実施例では、接続（ライン）は、ATMバーチャル・サーキット・サービスにおけるスイッチ・バーチャル・サーキットと定義されるが、ATMバーチャル・バス・サービスにおけるバーチャル・サーキットと、或いはATMローカル・エリア・ネットワークのエミュレーション・サービスにおける1組のMAC（Medium Access Control）アドレス（MACソース・アドレス、MACターゲット・アドレス）と定義することもできる。各ホップは、その入口にベスト・エフォート・トラフィック専用の入力バッファ521（IBBとす）、その出口にベスト・エフォート・トラフィック専用の出力バッファ（出力ポートPj）を持つ。ホップ50、51、52はUNIインタフェースからNNIインタフェースによって接続される。ホップ52に対して受信されたデータ・セルはキューイング・エリアの間に置かれ、3つのホップの間で確立された接続当たりキュー5211、...、521nが生じる。フロー制御情報は、ある上流ホップのベスト・エフォート接続にか、全てのの上流ホップのベスト・エフォート・トラフィック全体に適用される。ホップ52は、キューイング・エリア501のいくつもの上流キュー5011、5012、キューイング・エリア511の5111、5112からデータを受信できる。その場合、1接続に適用されるフロー制御情報（選択的バックプレッシャ）は、対応するキューイング・エリアに出力され、ベスト・エフォート・トラフィック全体に適用されるフロー制御情報（グローバル・バックプレッシャ）は、ホップ52に対して通信元の上流ホップ50、51それぞれに送られる。1つのベスト・エフォート接続に適用されるフロー制御情報（選択的バックプレッシャ）の場合の調整されない。例えばキュー5211がホップ52の選択的しきい値を超える場合、選択的ストップ・バックプレッシャ接続5211-5112がリンク524で送られ、ホップ51は、選択的バックプレッシャ情報を受信したときキュー5112のスケジューリングをストップする。ベスト・エフォート・トラフィック全体に適用されるフロー制御情報（グローバル・バックプレッシャ）の場合は、調整されるトラフィックは、関係ホップの間で転送されるベスト・エフォート・トラフィックの全てになる。例えばキューイング・エリア521がホップ52のグローバルしきい値を超えた場合、グローバル・ストップ・バックプレッシャ・リンク524、525が前記リンクを送られる。ホップ50は、グローバル・バックプレッシャ情報を受信したときキューイング・エリア501の全ての接続のスケジューリングをストップするが、キューイング・エリア502はフローを維持することができる。ホップ51は、グ

8
 ローバル・バックプレッシャ情報を受信したときキューイング・エリア511の全ての接続のスケジューリングをストップするが、キューイング・エリア512はフローを維持することができる。

【0015】図6は、ホップ入力バッファ構造及びネットワークの任意のホップによって受信/転送されるデータ・フローと制御フローを示す。フロー制御は、キューイング・エリアから対応する上流ノードに転送され、キューイング方式によって設定される様々なしきい値としきい値64、65、66、67によってトリガされる。IBB当たり上下のグローバルしきい値64、65がそれぞれ1つ、IBB当たり上下の選択的しきい値66、67がそれぞれ1つある。GBXPビット61は、IBB60に入るリンクにグローバル・バックプレッシャがかかっているかどうかを示す。カウント手段62は、前記IBBのアクティブ接続数（NAC）を表わし、カウント手段63は前記IBBの異なるキューに置かれたセルの合計を表わし、カウント手段611はキュー610に関係づけられて、前記キューに置かれたセルの数をカウントする。CBPXビット612は、接続610にバックプレッシャがかかっているかどうかを示し、カウント手段613（CBPXCNT）は、接続610がバックプレッシャ情報を送ってから前記接続によって受信されたセル数を表わす。SDビット614は、前記接続610によって転送されたセルを棄却するかどうかを示す。前記IBB60に入るリンク620それぞれに、指示手段6201、カウント手段6202、PDビット6203が関連づけられる。6201は前記リンク620の待ち時間を表わす。前記待ち時間はこの説明の後で定義する。カウント手段6202は、前記リンク620がグローバル・バックプレッシャ情報を送ってから前記リンクに対応する異なるキューに置かれたセル数を表わす。PDビット6203は、前記リンクによって転送されたセルを棄却するかどうかを示す。

【0016】図7は、ホップ出力バッファ構造及び受信/転送されるデータ・フローと制御フローを示す。ベスト・エフォート・サービスに対する出力バッファは1つあり、出力リンク727、725当たりEBB71、72である。フロー制御は下流ノードから受信され、ラウンド・ロビン・スケジューラ（RRS）68、69の適切な処置をトリガする。キュー721に入る接続が下流ホップによって選択的バックプレッシャを受けていると、関連する接続バックプレッシャ・ビットCBP721は1にセットされ、前記キュー721はラウンド・ロビン・スケジューラから追加削除される。リンク725が下流ホップによってグローバル・バックプレッシャを受けているとき、関連するグローバル・バックプレッシャ・ビットGBP726は1にセットされ、前記リンク725に関連するラウンド・ロビン・スケジューラ68はストップする。

9

【0017】図8は、グローバル・ストップ・バックプレッシャを生成する操作の流れ図、図9はグローバル・スタート/ストップ・バックプレッシャを受信する操作の流れ図である。ここでグローバル・ストップ処理について図6、図7、図8、図9を参照して説明する。セルがホップ50のIBB60にポートP1から受信されると、IBB60にキューイングされたセルの数63がIBB上グローバルしきい値64 (IBB_Th) に等しい場合、また前記ホップに入る上流ホップがまだバックプレッシャを受けていない場合 (GPX61が0)、ホップ50は、全ての入力ポートP1で後向きに、「グローバル・ストップ」という情報を持つグローバル・バックプレッシャを送り、GPXビットを1にセットする。図9で説明しているが、これを受信する上流ホップは、「グローバル・スタート」情報を受信するまでは、そのEBB (P1) からのセルのスケジューリングをストップする。これは、ラウンド・ロビン・スケジューラ68を停止させることによって行われる。情報伝送の往復で遅延が生じるため、余分なセルが受信されることがある。基準ホップ50 (バックプレッシャ信号を送る

ホップ) は、前記ポートP1から余分なセルを受信したときポートP1のPBPCNTカウンタを増分する。あるリンクの待ち時間の値は、あるホップと前記リンク端の上流ホップの間のセルの往復遅延の間に前記ホップのキューイング・エリアに置かれたセル数である。PBPCNTカウンタがリンク待ち時間値LATに達したとき、基準ホップは、前記リンクのポート棄却ビット (PD) をセットし、セルを棄却してグローバル・ストップ・バックプレッシャ情報を再送する。前記リンクからの他の全てのセルは、前記リンクのポート棄却ビットがリセットされるまで棄却される。入力セルの棄却に代わる方法は、PBPCNTカウンタがLAT値に達したときに、前記セルにタグをつけて受入れ、後に重大な輻射が生じた場合にそれらを棄却することである。待ち時間値をセットする場合には、上流ホップにおけるバックプレッシャ情報の処理に必要な時間を考慮して、往復遅延に安全余裕 (safety margin) を追加することもできる。

【0018】図10は、グローバル・スタート・バックプレッシャを生成する操作の流れ図である。ここでグローバル・スタート処理について図6、図7、図9、図10を参照して説明する。セルがホップ50のIBB60から転送されると、IBB60にキューイングされたセル数63が、IBBの下グローバルしきい値65 (IBB_Tl) に等しい場合、また前記ホップ50に入る上流ホップがすでにバックプレッシャを受けている場合 (ビットGPXが1)、ホップ50は後向きに、全ての

10

上流ホップをリセットする。図9で説明しているように、上流ホップはそれぞれ、そのEBB (P1) からセルのスケジューリングを再開する。ここでP1は基準ホップにポートで接続されている。これはラウンド・ロビン・スケジューラをリスタートさせることによって行われる。キューがラウンド・ロビン・スケジューラに追加されていない、選択的バックプレッシャがなかった接続からのトラフィックはリスタートしない。

【0019】ホップのキューイング・エリアのグローバルしきい値を定義する際、考慮すべき点は2つある。第1点は、前記ホップの上グローバルしきい値に関係する。損失のない動作を保証するには、しきい値より上の余分なキューイング・エリアが、前記ホップに接続された上流ホップの待ち時間合計より大きくならないようにしきい値を設定しなければならない。これは下記のように表わせる。

【数1】 $IBB_Th < IBB_Tl : f_{s, from} < \text{上流ホップ} > (\text{待ち時間} + \text{マージン}) : ef$

【0020】第2点は、前記ホップの下グローバルしきい値に関係し、このしきい値はアンダーラン動作がないように充分高く設定しなければならない。上グローバルしきい値IBB_Tlは、下グローバルしきい値より下の余分なキューイング・エリアが、前記ホップの上流ホップの最少待ち時間または往復遅延より大きくならないようにセットしなければならない。これは下記のように表わせる。

【数2】 $IBB_Tl > \text{最少 (待ち時間)}$

【0021】しかし、下グローバルしきい値が上グローバルしきい値に近すぎると、しきい値が多数重なるしまい制御情報のオーバーヘッドにつながる。本発明の好適な実施例に従って、下グローバルしきい値は上グローバルしきい値の1/2と定義されている。これは非アンダーラン条件を満足すると共に、トラフィック制御のオーバーヘッドとトラフィックのアンダーランのトレード・オフとして許容できる。

【0022】図6、図7に示す通り、ホップのIBBには2つの選択的しきい値が定義されている。選択的しきい値66 (IBB_Th) と選択的しきい値67 (IBB_Tl) は、前記IBBの全てのアクティブ接続に共通である。

【0023】図11は選択的ストップ・バックプレッシャを生成する操作の流れ図、図12は選択的スタート/ストップ・バックプレッシャを受信する操作の流れ図である。選択的ストップ処理について図6、図7、図11、図12を参照して説明する。キュー610に入る接続Ckからホップ50のIBB60でセルが受信され、IBBにキューイングされたポートP1のキュー610からのセル数611がIBB_Th66に等しいかより大きく、接続Ckがまだバックプレッシャを受けていない場合、ホップは後向きにポートP1で、「接続

50

C kの選択的ストップ」という情報を持つ選択的バックプレッシャを送る。図12で説明するが、これを受信するホップは接続C kの選択的スタート情報を受信するまでは、そのEBB(Pi)からの、接続C kからのセルのスケジューリングをストップする。これはラウンド・ロビン・スケジューラから接続のキューを追加解除することによって行なわれる。バックプレッシャ信号を送ったホップは、接続制御ブロックの、接続にバックプレッシャがかかった転送済みビットCBPXを1にセットする。これにより接続CBPX CNTのセル・カウンタがイネーブルになる。バックプレッシャがかかった接続から受信されたセルは全て、選択的スタートが送られるまでカウントされる。カウンタが所定接続C kからなるリンクの待ち時間である許容最大値LATに達した場合には、接続制御ブロックの選択的乗取ビットSDがセットされ、再びバックプレッシャ信号が送られる。好適な実施例では、SDビットがセットされたとき接続C kから受信された全てのセルが乗取られる。図13は、選択的スタート・バックプレッシャを生成する動作の流れである。ここでグローバル・スタートの処理について図6、図7、図11、図12を参照して説明する。接続C kからのセルがIBBから転送され、IBBにキューイングされたポートPiの接続C kからのセル数が、IBB-ST1に等しいかより少なく、接続C kが選択的バックプレッシャを受けている場合、ホップは後向きにポートPiで、接続C kの選択的スタートという情報を持つ選択的バックプレッシャを送る。これを受信するホップは、前記接続C kからなるリンクがグローバル・バックプレッシャを受けていない場合は、そのEBBから接続C kからのセルのスケジューリングを再開する。これはラウンド・ロビン・スケジューラの接続C kのキューを追加解除することによって行なわれる。バックプレッシャ信号を送ったホップは、CBPXビットとSDビット及び接続C kのCBPX CNTカウンタを0にリセットする。

【0024】図14は、選択的しきい値調整の表を示す。キューイング・エリアの上の選択的しきい値を定義する際にはいくつか制限を考慮する必要がある。値は提供されるサービスがフェアであり、アンダーラン条件がイネーブルにならず、制御トラフィックが可能な限り減少するように選ばれている。フェアネスの問題から、理論的には、選択的しきい値は、キューイング・エリア・サイズをアクティブ接続数NACで割った結果でなければならない。これは下記のように表わされる。

【数3】 $IBB_ST = IBB / NAC$

【0025】このような理論的關係を実現するのは難しい。この除算の結果が整数になることはほとんどないからである。逆に、図14に示したしきい値調整の原則は実現しやすく、先に述べた全ての条件を満足する。このしきい値調整原則についてはいくつかコメントを付ける

必要がある。

【0026】アクティブ接続数がNACの範囲の値の最低値、例えば範囲5乃至8の5である場合、NACにIBB-ST hををかけて得られる結果は、バッファのサイズより小さい。これは一見、ネットワーク・リソースの使用効率が悪いと解釈できる。しかし、第1点はバッファのうち用いられない部分、例えば $(IBB - (NAC * IBB_ST h))$ は、アクティブ接続数が増加すると減少する。実際、アクティブ接続が1つの場合には使用率は50%で、アクティブ接続が31の場合には97%、アクティブ接続数が32を超えると100%である。第2点は、アクティブ接続がほとんどない場合には、選択的しきい値は、接続が不利な状態にならない程度に充分に大きい。例えばキューイング・エリアが2048のセル・バッファとすると、1つのアクティブ接続は1024のセル・バッファ $(IBB / 2)$ を使用できる。このエリアが接続トラフィックで埋まる場合、上流の接続パスは下流の接続パスよりも飽和しにくい。ホップの役割はそととき、入ってくる上流トラフィックを下流の可能性に合わせるために制御することである。この観点からは、接続に大きいキューイング・エリアを使用することを認める必要はなく、1024のセル・バッファは大きい値とさえ思われる。しかしこの値により、ホップは上流のポートの適合理化といふその役割をスムーズに演じることができる。第3点は提案した方式が機能する環境は高度に動的であり、アクティブ接続数は永続的に発生する。しきい値調整ルールによって生じる未使用バッファ・エリアは、新しいアクティブ接続のための空きバッファを残す。

【0027】選択的しきい値は、アクティブ接続が32を超えるとき展開せず、64バッファに等しいままである。これは一見するとフェアネスの問題と解釈できる。しかし、アクティブ接続数が増えるとき選択的しきい値を下げると、しきい値はかなり小さくなる。例えばキューイング・エリアが2048のセル・バッファで、アクティブ接続数が512の場合、フェアネスの定義をあてはめると、しきい値は4バッファになる。これは明らかに許容できない制御トラフィックのオーバーヘッドにつながる。ある接続についてキューイングされたセルの数はしばしば上下のしきい値に重なるからである。ここで挙げているしきい値調整ルールは制御トラフィックを最少にする。またしきい値は理論上のものではないが、キューイング・エリアはアクティブ接続間でうまく共有される。すなわち、アクティブ接続全てについて、キューイングされたセルの数がしきい値に等しい場合、アクティブ接続は32を超えることはなく、キューイングは理想的である。また、アクティブ接続数が大きい(例えば512)場合、キューがしきい値で一杯のアクティブ接続数は必ず制限される。しきい値で10接続の場合、残り502接続は、キューイングされたセルの

13

平均数が $(2048 - (10 \times 64)) / (512 - 10) = 2.8$ セルに等しい。しきい値に達した10接続は選択的にストップさせられるので、またラウンド・ロビン・スケジューラ68 (キューイング・エリアの外部にある) はフェアであるので、残り502接続は占有するバッファが次第に増える。これよりシステムは、接続当たり4バッファのその平準状態になる。これは1BB (入力バースト・エフォート・バッファ) のパーフェクト・シェア (perfect share) である。

【0028】図15は、本発明の好適な実施例に従ったATMフロー制御セルの一般形式を示す。バックプレッシャ情報を伝えるために私設ATMセルが用いられる。バースト・エフォート・サービスのフロー制御セルのトラフィックは、予約帯域幅サービスを利用し、よってそのQOSによるメリットがある (遅延が少なくセル損失比が低い)。遅延が少ないことは往復遅延の減少に寄与し、セル損失比が低いことは制御セルの損失発生を最少にする。このような制御セルは、最大で12セル・スロット毎に1つ送信できる。よってバースト・エフォートのフロー制御トラフィックの予約帯域幅はリンク帯域幅の8%である。フロー制御トラフィックは、12番目のセル・スロットがあったとき、送るバックプレッシャ情報がある場合は、予約帯域幅サービス・キューに待機される場合としても制御セルは転送されるという意味でTDM (時間遅延多重) ライクである。送る制御情報がない場合には、予約帯域幅サービスのセルを転送するためにセル・スロットが用いられるか、或いは予約帯域幅サービス・キューが空の場合には、バースト・エフォート・サービスのセルを送信するためにセル・スロットが用いられる。フロー制御セルは専用の接続識別子、VPI (バーチャル・バス識別子)、VCI (バーチャル・チャネル識別子)、及び専用PTI (ペイロード・タイプ識別子) タイプ、PTI=110、によって識別される。セル損失優先ビットCLPは0にセットされる。HEC (ヘッダ・エラー制御) フィールドは、セル・ヘッダのビット・エラーを訂正するために用いられる。セル・ペイロードは、バックプレッシャ信号に関する全ての情報を含む。選択的バックプレッシャ信号の場合、伝えられる情報はバックプレッシャ信号と関係した接続のバーチャル・チャネル識別子 (VPI-VCI) とバックプレッシャ・コマンド (選択的スタート/ストップ) である。本発明の好適な実施例では、情報はセル・ペイロードの4フィールド (以下、プレッシャ・スロットと呼ぶ) で定義される。このフィールドで、VPI-VCIは、ちょうどATMセルのヘッダの正規のVPI-VCIのように置かれる。GBPビットはセルにフローバースト・バックプレッシャ情報があるかどうかを示し、SBPビットはプレッシャ・スロットに選択的バックプレッシャ情報があるかどうかを示す。GSRビットは (また、SSRビットは、前記グローバル情報 (選

14

択的) 情報がスタート情報かストップ情報かを示す。

【0029】私設セルは、48バイトのペイロードで12のバックプレッシャ・スロットを転送できる。これにより1つのフロー制御セルで、1乃至12の選択的バックプレッシャ情報を転送でき、よってこの形態により、フロー制御トラフィックとこれに対する1対1の形態によって必要になる帯域幅が減少する。また、この形態はバックプレッシャ情報の転送を高速化する。送信されるバックプレッシャ情報が1より少ない場合、それらは制御セルで最初のプレッシャ・スロットを始点に連続したプレッシャ・スロットを使用する。グローバル・バックプレッシャ信号の場合、伝えられる情報はスタートかストップのコマンドである。本発明の好適な実施例では、グローバル・スタート/ストップ・バックプレッシャ・コマンドが私設セルの最初のプレッシャ・スロットに2ビットでコード化される。

【0030】まとめとして本発明の構成に関して以下の事項を開示する。

【0031】(1) 複数の接続からなるリンクによって相互接続された複数のホップを含む通信システムに実現され、前記接続内のトラフィックはバースト・エフォート・サービスと予約帯域幅サービスに共有され、各ホップにおいて、前記バースト・エフォート・サービスの制御は、単位接続とホップ毎のグローバル・バックプレッシャ・メカニズムの両方からなることを特徴とする、トラフィック制御装置。

(2) 前記ホップ毎のバックプレッシャ・メカニズムは、前記ホップに入るトラフィックが所定しきい値を超えたときに上流ホップにバックプレッシャ・プリミティブを生成する、前記(1)記載のトラフィック制御装置。

(3) 各ホップは、サイズNの入力バッファと、出力リンク当たりサイズMの出力バッファとからなり (MとNは1より大きい整数)、いずれのバッファもバースト・エフォート・トラフィック専用であって、前記ホップに接続された接続毎にキューを提供するキューイング・エリアとなり、前記接続は、その対応するキューに置かれたトラフィック・セルが少なくとも1つある場合はアクティブとみなされる、前記(1)または(2)に記載のトラフィック制御装置。

(4) 前記入力バッファは、グローバルしきい値IBB_{th}とグローバルしきい値IBB_{tl}よりなり、IBB_{th} > IBB_{tl}であって、前記しきい値は、前記しきい値を超える前記入力バッファ内の余分なバッファ・エリアが、上流ホップの往復遅延合計より大きくなるようにセットされ、前記下しきい値は、上流ホップの最少往復遅延よりも大きい、前記(3)記載のトラフィック制御装置。

(5) 前記入力バッファは、選択的上しきい値IBB_{sth}と選択の下しきい値IBB_{stl}よりなり、IBB_{sth}

50

15

$sth > IBbstl$ であって、前記しきい値は、上流ホップからのアクティブ接続の数に従って動的に調整される、前記(3)または(4)に記載のトラフィック制御装置。

(6) 各ホップにおいて、前記上下の選択的しきい値 $IBbstl$ 、 $IBbsth$ は、

前記ホップに入るアクティブ接続の数が1と2の間にある場合、 $IBbsth = N/2$ 、 $IBbstl = N/4$

前記ホップに入るアクティブ接続の数が3と4の間にある場合、 $IBbsth = N/4$ 、 $IBbstl = N/8$

前記ホップに入るアクティブ接続の数が5と8の間にある場合、 $IBbsth = N/8$ 、 $IBbstl = N/16$

前記ホップに入るアクティブ接続の数が9と16の間にある場合、 $IBbsth = N/16$ 、 $IBbstl = N/32$

前記ホップに入るアクティブ接続の数が17と32の間にある場合、 $IBbsth = N/32$ 、 $IBbstl = N/64$

前記ホップに入るアクティブ接続の数が33より大きい

か33に等しい場合、 $IBbsth = N/32$ 、 $IBbstl = N/64$

というルールに従って調整される、前記5記載のトラフィック制御装置。

(7) あるホップ接続のパッファ・エリアにキューイングされたセル数が、前記動的選択的しきい値を超えるとき、前記ホップは、前記接続からなるリンク上で後向きに、「前記接続の選択的ストップ」情報を持つ選択的バックプレッシャを送る、前記(2)乃至(6)に記載のトラフィック制御装置。

(8) ある接続のパッファ・エリアにキューイングされたセル数が、前記動的選択的しきい値に等しいかより小さく、前記接続が選択的バックプレッシャを受けているとき、前記ホップは、前記接続よりなるリンクで後向きに、「前記接続の選択的スタート」情報を持つ選択的バックプレッシャを送る、前記(2)乃至(7)に記載のトラフィック制御装置。

(9) あるホップのキューイング・エリアに置かれたセル数が、前記グローバルしきい値を超えるとき、前記ホップは、前記リンクをグローバルに停止させるために、前記ホップに入る上流リンクで後向きに、「前記入

力リンクのグローバル・ストップ」情報を持つグローバル・バックプレッシャを送る、前記(2)乃至(8)に記載のトラフィック制御装置。

(10) あるホップのキューイング・エリアに置かれたセル数が、前記グローバルしきい値に等しいかより小さく、前記ホップに入るリンクがすでにグローバル・バックプレッシャを受けている場合、前記ホップは、前記リンクをグローバルにリスタートさせるために、前記入

力リンクで後向きに「前記入力リンクのグローバル・ス

タート」情報を持つグローバル・バックプレッシャを送る、前記(2)乃至(9)に記載のトラフィック制御装置。

(11) 選択的に停止された接続は、前記接続が選択的スタート・バックプレッシャを受け取り、前記接続を含む前記リンクがグローバル・バックプレッシャを受けていない場合にのみリスタートでき、グローバルに停止されたリンクは、前記リンクがグローバル・スタート・バックプレッシャを受け取った場合にのみリスタートできる、前記(2)乃至(10)に記載のトラフィック制御装置。

(12) ホップは、前記ホップに入る各リンクに関連づけられ、前記ホップが前記入力リンクにグローバル・ストップ・バックプレッシャを送った後に前記入力リンクから受信した余分なセルをカウントする手段を含み、前記入力リンクのカウント値が前記入力リンクの往復遅延より大きい前記遅延に等しいとき、前記入力リンクからの余分なセルは棄却され、前記ホップは前記グローバル・ストップ・バックプレッシャを前記入力リンクに再送する、前記(1)乃至(11)に記載のトラフィック制御装置。

(13) ホップは、前記ホップに入る各接続に関連づけられ、前記ホップが前記接続よりなるリンクで前記接続に選択的ストップ・バックプレッシャを送った後に前記接続から受信した余分なセルをカウントする手段を含み、前記接続のカウント値が、前記接続よりなる前記リンクの往復遅延より大きい前記遅延に等しいとき、前記接続からの余分なセルは棄却され、前記ホップは前記選択的ストップ・バックプレッシャを前記接続に再送する、前記(1)乃至(12)に記載のトラフィック制御装置。

(14) 予約帯域幅の1部がベスト・エフォート・サービスの制御トラフィックに予約された、前記(1)乃至(13)に記載のトラフィック制御装置。

(15) 前記バックプレッシャ情報はATM制御セルによって伝達され、前記セルは、前記ベスト・エフォート・サービスの前記制御トラフィックに予約された前記予約帯域幅の前記1部によって転送される、前記(1)乃至(14)に記載のトラフィック制御装置。

(16) 前記ATM制御セルは、1乃至12の選択的バックプレッシャ情報と1つのグローバル・バックプレッシャ情報を伝達できる、前記(1)乃至(15)に記載のトラフィック制御装置。

(17) 接続は、ATMバーチャル・サーキット・サービスのスイッチ・バーチャル・サーキットと定義されるか、ATMバーチャル・バス・サービスのバーチャル・サーキットと定義されるか、或いはATMローカル・エリア・ネットワークのエミュレーション・サービスにおける1組のメディア・アクセス制御(MAC)アドレス(MACソース・アドレス、MACターゲット・アドレ

10

20

30

40

50

ス)と定義される、前記(1)乃至(16)に記載のトラフィック制御装置。

【0032】

【発明の効果】以上説明したように本発明によればフェアネスを実現しつつ、バッファ容量が小さくても損失のないトラフィック制御を可能とする。また、輻輳したリンクや接続でトラフィック制御情報がブロックされることもなく、出力するトラフィック制御も少ないトラフィック制御装置を提供できる。

【図面の簡単な説明】

【図1】従来の予約帯域幅サービスの図である。

【図2】従来の非予約帯域幅サービスの図である。

【図3】複数のホップからなるネットワークの図である。

【図4】6ホップ構成の概略図である。

【図5】本発明の好適な実施例に従ったホップ毎のフロー制御装置の詳細図である。

【図6】ホップ入力バッファの構造図である。

【図7】ホップ出力バッファの構造図である。

【図8】グローバル・ストップ・バックプレッシャを生成する動作の流れ図である。

【図9】グローバル・スタート/ストップ・バックプレッシャを受信する動作の流れ図である。

【図10】グローバル・スタート・バックプレッシャを生成する動作の流れ図である。

【図11】選択的ストップ・バックプレッシャを生成する動作の流れ図である。

【図12】選択的スタート/ストップ・バックプレッシャ

を受信する動作の流れ図である。

【図13】選択的スタート・バックプレッシャを生成する動作の流れ図である。

【図14】選択的しきい値調整の表を示す図である。

【図15】ATMフロー制御セルを示す図である。

【符号の説明】

50、51、52 ホップ

60 IBB

61 GBPXビット

62、63、611、613、6202 カウント手段

64、65 グローバルしきい値

66、67 選択的しきい値

68、69 ラウンド・ロビン・スケジューラ

71、72 EBB

501、502、511、512 出力バッハ

521 入力バッハ(1BB)

524、525 バックプレッシャ・リンク

610、721 キュー

612 CBPXビット

614 SDビット

620 リンク

725、727 出力リンク

726 GBPR

5011、5012 上流キュー

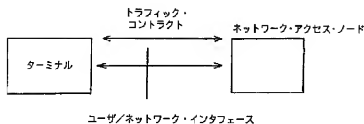
5111、5112、5211 キューイング・エリア

6201 指示手段

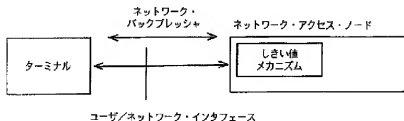
6203 PDビット

7211 CBPR

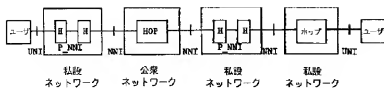
【図1】



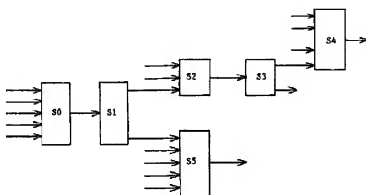
【図2】



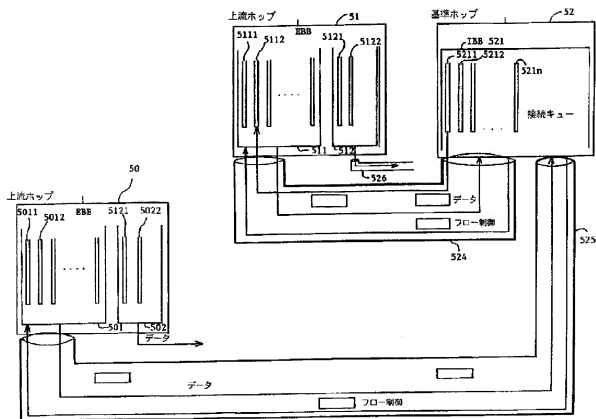
【図3】



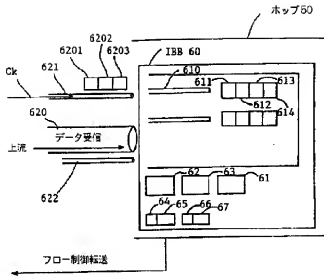
【図4】



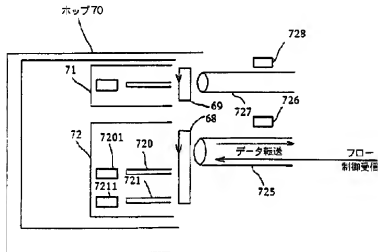
【図5】



【図6】



【図7】

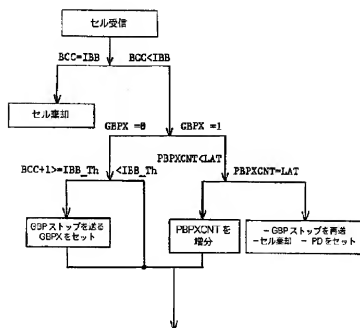


【図14】

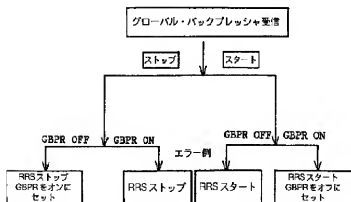
しきい値調整の原則

キュー数	上しきい値	下しきい値
1 - 2	136/2セル	133/4セル
3 - 4	130/4セル	129/8セル
5 - 8	108/8セル	100/16セル
9 - 16	108/16セル	100/32セル
17 - 32	108/32セル	100/64セル
32より大きい	108/32セル	100/64セル

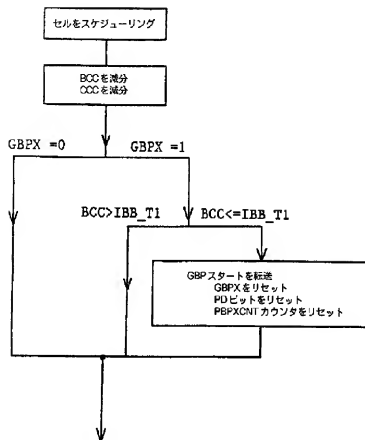
【図8】



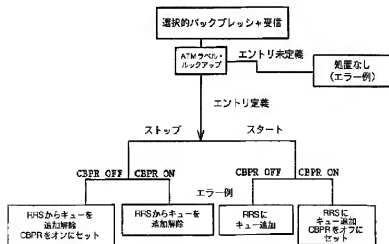
【図9】



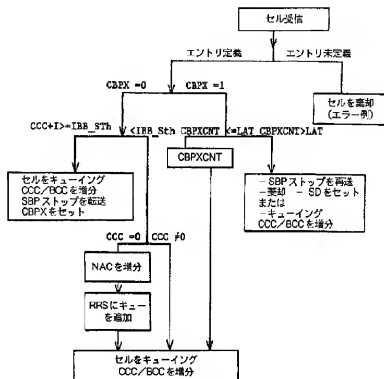
【図10】



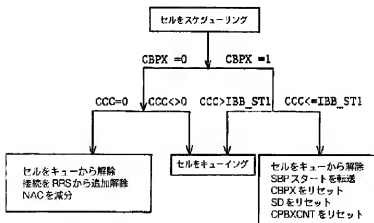
【図12】



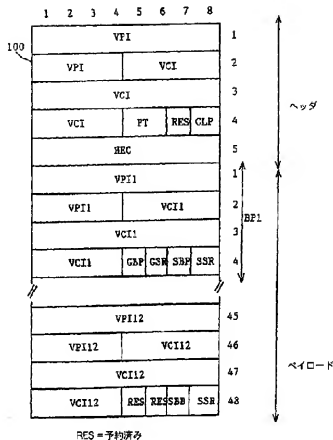
【図11】



【図13】



【図15】



フロントページの続き

(72)発明者 ファブリス・パーブランケン
フランス06610、ラ・ゴーズ、ルート・
デ・セント・ローレント 1706

(72)発明者 ダニエル・オーサッティ
フランス06800、カグネス・サーマー、
アリー・デス・ボレックス 11、レ・ド
ーフィン・ブル

(72)発明者 クラウド・バツソ
フランス06200、ニース、アベニュー・
デ・ラ・ラターネ 252、ラ・ラターネ・
デス・アングス、エイ 3